

```

1 GCGCTCGCCG CGCGGGGAGA AAGAAGCCGC GCCCAGCCCG GCGTCCCGAG
51 CAGCGCAGGG GAGGATCCCC GCGCAGTGAC CCGGGAGCCA CCACAGACTC
101 TGGGAGGCTC GCGGGCTGGA GCAGCAGGCA GCTCCCCGCA GCTCCCGCG
151 CTTCCAGGCA GCTCTCTGAG CCGTGCCAGA GGCCCCGGCC GCCATTCCCA
201 GCCCCGAGCC ATGATGAAGA CTTTGTCCAG CGGGAAGTGC ACGCTCAGTG
251 TGCCCGCCAA AAATCATA CCGATGGTGG TGCTGGGTGC CTCTCGGGTG
301 GGCAAGAGCT CCATCGTGTC TCGCTTCCTC AATGGCCGCT TTGAGGACCA
351 GTACACACCC ACCATCGAGG ACTTCCACCG TAAGGTATAC AACATCCGCG
401 GCGACATGTA CCAGCTCGAC ATCCTGGATA CCTCTGGCAA CCACCCCTTC
451 CCCGCCATGC GCAGGCTGTC CATCCTCACA GGGGATGTCT TCATCCTGGT
501 GTTCAGCCTG GATAACGGG AGTCCTTCGA TGAGGTCAAG CGCCTTCAGA
551 AGCAGATCCT CGAGGTCAAG TCCTGCCTGA AGAACAAGAC CAAGGAGGCG
601 GCGGAGCTGC CCATGGTCAT CTGTGGCAAC AAGAACGACC ACGGCGAGCT
651 GTGCCGCCAG GTGCCACCA CCGAGGCCGA GCTGCTGGTG TCGGGCGACG
701 AGAACTCCGC CTAATTCGAG GTGTGCGCCA AGAAGAACAC CAACGTGGAC
751 GAGATGTTCT ACGTGCTCTT CAGCATGGCC AAGCTGCCAC ACGAGATGAG
801 CCCCGCCCTG CATCGCAAGA TCTCCGTGCA GTACGGTGAC GCCTTCCACC
851 CCAGGCCCTT CTGCATCGCG CGCGTCAAGG AGATGGACGC CTATGGCATG
901 GTCTCGCCCT TCGCCCGCCG CCCCAGCGTC AACAGTGACC TCAAGTACAT
951 CAAGGCCAAG GTCCTTCGGG AAGGCCAGGC CCGTGAGAGG GACAAGTGCA
1001 CCATCCAGTG AGCGAGGGAT GCTGGGGCGG GGCTTGGCCA GTGCCCTCAG
1051 GGAGGTGGCC CCAGATGCCC ACTGTGCGCA TCTCCCCACC GAGGCCCGCG
1101 CAGCAGTCTT GTTCACAGAC CTTAGGCACC AGACTGGAGG CCCCCGGGCG
1151 CTGGCCTCCG CACATTCGTC TGCCTTCTCA CAGCTTTCCT GAGTCCGCTT
1201 GTCCACAGCT CTTGGTGGT TTCATCTCCT CTGTGGGAGG ACACATCTCT
1251 GCAGCCTCAA GAGTTAGGCA GAGACTCAAG TTACACCTTC CTCTCCTGGG
1301 GTTGAAGAA ATGTTGATGC CAGAGGGGTG AGGATTGCTG CGTCATATGG
1351 AGCTCCTGG GACAAGCCTC AGGATGAAAA GGACACAGAA GGCCAGATGA
1401 GAAAGGTCTC CTCTCTCCTG GCATAACACC CAGCTTGGTT TGGGTGGCAG
1451 CTGGGAGAAC TTCTCTCCCA GCCCTGCAAC TCTTACGCTC TGGTTCAGCT
1501 GCCTCTGCAC CCCCTCCAC CCCCAGCACA CACACAAGTT GGCCCCCAGC
1551 TGCGCTGAC ATTGAGCCAG TGGACTCTGT GTCTGAAGGG GCGTGGCCA
1601 CACCTCCTAG ACCACGCCCA CCACTTAGAC CACGCCCACC TCCTGACCGC
1651 GTTCTCAGC CTCTCTCCT AGGTCCCTCC GCCCGACAGT TGTGCTTGT
1701 TGTGTTGCA GCTGTTTTCG TGTATGTAT AGTAGTAGAA ATGGAATCA
1751 TTGTACTGTA AAAGCCTAGT GACTCCCTCC TTGGCCAGGC CCTACCCAG
1801 TTCAGATCCA CGGCCTCCAC CCGGACGCC TTCTCTCTCT GCTCCCAAAC
1851 AGGGTTTCCG TGGCCTGTTT GCAGCTAGAC ATTGACCTCC GCCATTGAGC
1901 TCCACGGTTT ACAGACAATT GCACAAGCGT GGGGTGGGCA GGCCAGGACT
1951 GCTTTTTTTT AATGCTCCCA TTTCACAGAG GATACCACCG AGACTCGGAG
2001 GGGACACGAT GAGCACCAGG CCCCACCTTT GTCCCTTAGC AAATTCAGGG
2051 TACAGTCCA CCTAGAACCA GGCTGCCCTC TACTGTGCTC GTTCTCAAG
2101 CATTATTATA GCACCTACTG GGTGCTGGGT TCACTGTGTC CTAGGAAACC
2151 AAGAGGGTCC CCACTCTGG CCTCTGCCG CCCCTGCTGC CCCACCACCT
2201 TCTGCACACA CAGCGGTGGG GAGGCGGGGA GGAGCAGCTG GGACCCAGAA
2251 CTGAGCCTGG GAGGATCCG ACAGAAAAGC TCAGGGCGGG TCTTCTCCTT
2301 GTGCCCGGGA TTGGGCTATG CTGGGTACCA CCATGTACTC AGGCATGGTG
2351 GGTTTTGAAC CCATAAACCA AAGGCCCTTG TCATCAGCTC TTAACAAGTA
2401 TATTTTGTAT TTTAATCTCT CTAAACATAT TGAAGTTTTA GGGCCCTAAG
2451 GAACCTTAGT GATCTTCTAT TGGGTCTTTC TGAGGTTTCA AGAGGGTAAG
2501 TAACCTCCTC CAGGTCACAC AGCAAGTCTG TGGGTGGCAG AAGCAAGCTA
2551 GCGCTGGGCA TTCAGTACAT ACCACGATGT GCTCCCTCTC TTGATGCTTG
2601 GCCCCTGGGG CTTTCAGGGC TTTGGGACAT CTGTCTCTCA ACCCTCTCCC
2651 TAGATCAGTC TGTGAGGGTC CCTGTAGATA TTGTGTACAC CATGCCCATG
2701 TATATACAAG TACACACAGA GTACACACA GATGTACACA TGCTCCAGCC
2751 CCAGCTCTGC ATACCTGCAC CTGCACCCCA GCCTTGGCCC CTGCTGCGT
2801 CTGTGCTCAA AGCAGCAGCT CCAACCTGCT CTCTGTCCCC TTCCCCACCC
2851 ACTGCCTGAG CTTTCTGAGC AGACCAGGTA CCTTGGCTGC ACCGGTGTGT
2901 GGCCCGCTCT CACCCAGGCA CAGCCCCGCC ACCATGGATC TCCGTGTACA
2951 CTATCAATAA AAGTGGGTTT GTTACAAAAA AAAAAAAAAA AAAAAAAAAA
3001 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
3051 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA

```

FIGURE 1, page 1 of 3

FEATURES:

5'UTR: 1-210
 Start Codon: 211
 Stop Codon: 1009
 3'UTR: 1012

5' UTR ANALYSIS:

Query=cDNA clone
 Sbjct: genomic sequence

5' UTR Exon 1, non-coding

Score = 58.0 bits (29), Expect = 2e-10
 Identities = 29/29 (100%)
 Strand = Plus / Plus

Query: 1 ggcgtcgccgcgcggggagaaagaagccg 29
 |||||
 Sbjct: 1535 ggcgtcgccgcgcggggagaaagaagccg 1563

5' UTR Exon 2, non-coding

Score = 349 bits (176), Expect = 3e-98
 Identities = 176/176 (100%)
 Strand = Plus / Plus

Query: 26 gccgcgcccagcccggcgctcccgagcagcgcagggaggatccccgcgcagtgacccggg 85
 |||||
 Sbjct: 2001 gccgcgcccagcccggcgctcccgagcagcgcagggaggatccccgcgcagtgacccggg 2060

Query: 86 agccaccacagactctgggaggctcggcggtggagcagcagggcagctccccgcagctcc 145
 |||||
 Sbjct: 2061 agccaccacagactctgggaggctcggcggtggagcagcagggcagctccccgcagctcc 2120

Query: 146 cggcgcttccaggcagctctctgagccgtgccagaggcccggcccgcattcccag 201
 |||||
 Sbjct: 2121 cggcgcttccaggcagctctctgagccgtgccagaggcccggcccgcattcccag 2176

Score = 563 bits (284), Expect = e-162
 Identities = 284/284 (100%)
 Strand = Plus / Plus

5' UTR Exon 3, Protein coding region begins at nucleotide 211

Query: 199 cagccccgagccatgatgaagactttgtccagcggaactgcacgctcagtggtgcccgcc 258
 |||||
 Sbjct: 7474 cagccccgagccatgatgaagactttgtccagcggaactgcacgctcagtggtgcccgcc 7533

Homologous proteins:

Top BLAST Hits

	Score	E
CRA 18000005194969 /altid=gi 10047088 /def=ref NP_055125.1 sim...	538	e-152
CRA 18000005238449 /altid=gi 5059122 /def=gb AAD38928.1 AF13440...	512	e-144
CRA 332000009620725 /altid=gi 8118457 /def=gb AAF72997.1 AF2620...	342	2e-93
CRA 18000005232775 /altid=gi 7706359 /def=ref NP_057168.1 ras-...	342	2e-93
CRA 87000000006130 /altid=gi 7230768 /def=gb AAF43090.1 AF23915...	342	3e-93
CRA 18000005090459 /altid=gi 6677673 /def=ref NP_033052.1 RAS,...	341	6e-93
CRA 89000000197633 /altid=gi 7295299 /def=gb AAF50620.1 (AE003...	228	8e-59
CRA 105000014645240 /altid=gi 10503969 /def=gb AAG17979.1 AF177...	214	1e-54

1 MMKTLSSGNC TLSVPAKNSY RMVVLGASRV GKSSIVSRFL NGRFEDQYTP
 51 TIEDFHRKVY NIRGDMYQLD ILDTSGNHFF PAMRRLSILT GDVFLVFSL
 101 DNRESFDEVK RLQKQILEVK SCLKNKTKEA AELPMVICGN KNDHGELCRQ
 151 VPTTEAELLV SGDENSAYFE VSAKKNTNVD EMFYVLFSMA KLPHEMSPAL
 201 HRKISVQYGD AFHPRPFCMR RVKEMDAYGM VSPFARRPSV NSDLKYIKAK
 251 VLREGQARER DKCTIQ

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION

N-glycosylation site

Number of matches: 2

1 9-12 NCTL
 2 125-128 NKTK

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE

cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 4

1 84-87 RRLS
 2 174-177 KKNT
 3 202-205 RKIS
 4 236-239 RRPS

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE

Protein kinase C phosphorylation site

Number of matches: 2

1 19-21 SYR
 2 172-174 SAK

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE

Casein kinase II phosphorylation site

Number of matches: 5

1 51-54 TIED
 2 105-108 SFDE
 3 154-157 TEAE
 4 161-164 SGDE
 5 177-180 TNVD

[5] PDOC00007 PS00007 TYR_PHOSPHO_SITE

Tyrosine kinase phosphorylation site

220-228 RRVKEMDAY

[6] PDOC00266 PS00294 PRENYLATION

Prenyl group binding site (CAAX box)

263-266 CTIQ

[7] PDOC00016 PS00016 RGD

Cell attachment sequence

63-65 RGD

[8] PDOC00017 PS00017 ATP_GTP_A

ATP/GTP-binding site motif A (P-loop)

26-33 GASRVGKS

BLAST Alignment to Top Hit:

>CRA|18000005194969 /altid=gi|10047088 /def=ref|NP_055125.1| similar
to mouse Ras, dexamethasone-induced 1; tumor endothelial
marker 2 [Homo sapiens] /org=Homo sapiens /taxon=9606
/dataset=nraa /length=278
Length = 278

Score = 538 bits (1372), Expect = e-152
Identities = 265/266 (99%), Positives = 265/266 (99%)

Query: 1 MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY 60
MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY
Sbjct: 13 MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY 72

Query: 61 NIRGDMYQLDILDTSGNHPPFAMRRLSILTGDVFIILVFSLDNRESFDEVKRLQKQILEVK 120
NIRGDMYQLDILDTSGNHPPFAMRRLSILTGDVFIILVFSLDNRESFDEVKRLQKQILEVK
Sbjct: 73 NIRGDMYQLDILDTSGNHPPFAMRRLSILTGDVFIILVFSLDNRESFDEVKRLQKQILEVK 132

Query: 121 SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENSAYFEVS AKKNTNVD 180
SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDEN AYFEVS AKKNTNVD
Sbjct: 133 SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENCAYFEVS AKKNTNVD 192

Query: 181 EMFYVLFMAKLPHEMSPALHRKISVQYGDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV 240
EMFYVLFMAKLPHEMSPALHRKISVQYGDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV
Sbjct: 193 EMFYVLFMAKLPHEMSPALHRKISVQYGDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV 252

Query: 241 NSDLKYIKAKVLREGQARERDKCTIQ 266
NSDLKYIKAKVLREGQARERDKCTIQ
Sbjct: 253 NSDLKYIKAKVLREGQARERDKCTIQ 278

Complete Amino Acid Sequence of gi|10047088 /def=ref|NP_055125.1| similar
to mouse Ras, dexamethasone-induced 1; tumor endothelial
marker 2 [Homo sapiens] /org=Homo sapiens /taxon=9606
/dataset=nraa /length=278
Length = 278

NOTE: UNDERLINED RESIDUES NOT PRESENT IN THE PROTEIN OF THE PRESENT INVENTION

1 mpaslallqp rammktlssg nctlsvpakn symmvvlgas rvgkssivsr flngrfedqy
61 tptiedfhrk vynirgdmq ldildtsgnh pfpamrrlsi ltgdvfilvf sldnresfde
121 vkrlqkqile vksclnkntk eaaelpmvi gnkndhgelt rqvptteael lvsgdencay
181 fevsakkntn vdemfyvlfs maklphemp alhrkisivqy gdafhprpfc mrrvkemday
241 gmvspfarrp svnsdlkyik akvlregqar erdkctiq

>CRA|18000005238449 /altid=gi|5059122 /def=gb|AAD38928.1|AF134409_1
(AF134409) Rhes protein [Rattus norvegicus] /org=Rattus
norvegicus /taxon=10116 /dataset=nraa /length=266
Length = 266

Score = 512 bits (1304), Expect = e-144
Identities = 252/266 (94%), Positives = 256/266 (95%)

Query: 1 MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY 60
MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY
Sbjct: 1 MMKTLSSGNCTLSVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYTPTIEDFHRKVY 60

Query: 61 NIRGDMYQLDILDTSGNHPPFAMRRLSILTGDVFIILVFSLDNRESFDEVKRLQKQILEVK 120
NI GDMYQLDILDTSGNHPPFAMRRLSILTGDVFIILVFSLD+RESFDEVKRLQKQILEVK
Sbjct: 61 NIHGDMYQLDILDTSGNHPPFAMRRLSILTGDVFIILVFSLDSRESFDEVKRLQKQILEVK 120

Query: 121 SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENSAYFEVS AKKNTNVD 180
SCLKNKTKEAAELPMVICGNKNDH ELCRQVP EAELLVSGDEN AYFEVS AKKNTNV+
Sbjct: 121 SCLKNKTKEAAELPMVICGNKNDHSELCRQVPAMEAELLVSGDENCAYFEVS AKKNTNVN 180


```

1 CTCTCTGACT CTTTGCCTCC TCTCTGACTC CCTGCCTCCT CTCTCTGTCT
51 CCCTGCCTCC TCTGTCTGAC TCCCTGCCTC CCCTCTCTGT CTCCTGCCT
101 CCTCTCTCTG ACTCTCTGCC TCCTCTCTCT GACTCCCTGC CTCCTCTCTC
151 TGATTCCCTG CCTCTTTGAC CCTCTGCCTC CTCTCTTTGA CTCCTGCCT
201 CCTCTCTCCG ATTCTCTGCG TCTTTGACTC CCTGCCTCCT CTCTCTGACT
251 CCCTGAAGCT CATTCACTCA TTGCTATCAA CTCGTCTGTA CCAAGCTCTA
301 GGCTGGAGGC TGGGCAGGGC AATGATGGAG ACAAATACTG TCCCTGGGAG
351 CTTCTGGCCC CTTTCCCATC CTGTTTAGAC AGAAGTGACC GCCAGCAGAG
401 TCAAGCTGTC TGCAGAAGGA CTTGGGGAGG GGGCTGTCAT GGGGTAGGGC
451 TTCTTTCCCC CCATCTCTGC TGAAGGCCA GGCTGGCTGA GACAGCCCCG
501 GCAGAGACTG AGAAGGGCTC CCTGCTGTGG TCTGGCAGCC CCCTCTCCAC
551 CCTCTCTCTC CTCATTTTCT GCCTCCCACA CGTATGCCCT GGGCACCTCA
601 TCAGGGCTGC CCTAGGGGAG GGCCCTCCTT GGCACAGCCC CTGGGCCAGT
651 CAGGTGGTTG AGGCTGAGGA GAGAAGGTCC CAGAGTGGGG CTTCAGGCAA
701 ACCCAAAGAC AGAGCCCTTT GCCATTTGAT GAATGCACAG ACCCTTTATT
751 GAGCCCCTGC TCTGTTTATG GCATGGCAGT TTTGTGGGAT AAATCAAAG
801 ACAGCTTTAG GTGGGAGCTG GGTGGGGGAT GTGGGGGTCT TAGGCTTGAA
851 CTACTACCCA GCCTCCTTTG TTAACCAAGT AGCTAGTCAC GTAGCCTTCT
901 GAGCTCGGGG CAGACCACCT GGGATCAAAC CTCTCCTCTG CTGGTACTG
951 GCTGTGCAAC TGTAAGCAAG TAATTTAACC TCTCTGTGCC TCAGTTTCCT
1001 CATCTGTAAG TTGGAGAATA ACACCACCTG CTTTCTGGGG TTATGAAGGG
1051 AGAAATAGGT TAACATGTGT GCAGCACTTA GAACACTCTG GCATATTTTA
1101 GCTGCAAAAT GAATGCCAGC TATGATTATT TCTATACTTA GTGCGGGGCT
1151 TGGACACTCG CATGGGCTCA AGTGGCAGCA GTTGTCTGTC TTGTGGCTCC
1201 AGGCTGGGGG TCCGCCGTGT GCTGAGCTGG CTTATTGTGC ACGTCCCTTT
1251 GTGATTCATT CATCGAAGTC ACATTAGTAG CTTAGAAGTG ACCGTAGTGG
1301 GAGCATTTAC GCCATGGAAA TTGGCAATAG GGCTTTTAAC AAAGGTATTT
1351 TTGAGAGCCG GTTTCCTGCA CAGAGGCTGG TAGTTGGGCA GGGTGAGCAG
1401 ATCCAGATGT GTGCCAGGGA CTCGCACGCA GGCAATCTCT CCACCTCCAG
1451 TGGCCATCTC AGACCTTAGC TTCATGATAG CCAGGAAGCG ATGGTGTGG
1501 AAAGCGCCTT GGGTCAATGG GCGAGGCACT CAAGGAAACC GACTTGGGGC
1551 ATCTGGGGT GGGGACCGAG TTTGGGCACA TACAGCCCTT TGTGTGAATT
1601 TAAAAACAGT GCCTTTTCTT CTACACAAGA TGCCCTTTCT TCTGGGATAC
1651 AGCCCCCACC TCTGGGATGC AGCCCCCACT TGCCCAACCA GCCATGCGCC
1701 TTGTGCAGTA TCCAACCTGC ACAACCTGTG GCAGCCTGTG GAAGACCGAG
1751 GGGATTGATA TTTCAGCAGG CCTGTGCCCC TTTGCAGTTC AGGGGCTGGA
1801 AAGTCTCCTT CTGGAGAGGG GAGGGATTCC TGCAAGGGTG AGGAGATCAG
1851 AGAGGCCTTC AGAGAGCAGG TGGCACTTGA GCCAGACCCT GAAACATAAG
1901 GGGAAGAGGG GTTCTGTCAG AGGGGTGGCA TGAGCAAAGG AGTGGAGGCT
1951 GATCTCAGCA GAGCTCAAAC TGACGAGGGT GACTGGGGTC AGGGGTCTG
2001 GGGCGGGGAT TCTGGTGGGC GCTAAGGTAG GAAAGGAGGG AGGGCTGGGC
2051 TGTGAAGAGC CTTTGGGGTG AGCCTGGTGG AGCCTGCGGG TTTGCTTATA
2101 CAAGAGCTTG GATCCATGTC GGCTCTTTT ATGAGGTCAA GAGGCTCCCA
2151 TAGAAAAGCTC TGAGTTTGCC CCAGAACCAT AACCTTGGA GATGGGAGGG
2201 AAGCTTGAGC CAGCCATGGG TCGTTCCCCA TTCCACATCC TCTACTCCGG
2251 GCCTCTGGGT CTCCTGGAGG CAAGTAAACA CCTAGGGCCT GGGAGGCAAA
2301 AATATCCGGG CAGGTCATGG AGCGGAGGGA GCCCGCCAGA TGCAGAGCAC
2351 AGGTCTAAAG GTGGGTCTCT CTGAGGTGGC TGCAAGGAGCA ACCCAGGCA
2401 TTGGGCTTGG AGCATGCGGT GTGGACATAG CCTTCCCTTC TTCCCAGGAG
2451 GGCTGAATGG CCACAGAACC ACCCCTGCC CCAGGCTTAA GAAATGCATG
2501 CTAGTGCCCT CCCCATGTCT TATCCTAGAA TCACAGGCTC CGGGAAAGCC
2551 AGATGGATGA ACCAGGGAAG GAACGGATTG TCACCATAGA TACCATTTT
2601 GAGATTTTAC CATGTGCTGA GCCCTTTGCA ACAACTCTAT GAATTGGGCT
2651 CATTTTGTCAG ATGAGAAAAG TGAATTCTAG AGAGGTAAAG CTACTAGCCC
2701 AAGATCAGTA GCTAGAGGCA AGGCAAGGAT TCAAATCCCA GGAGTCCGGT
2751 GCTTGCAATA ATGAAAGGAT GAATGAACGG ATATTGAGTG AGTGAGTGA
2801 TGAAGGAAGG AGTAAAGGAG AGGGCATGAA TGAATGAGAG GGTAGAACTC
2851 CAAGACCCCT TAGAACCTCG TCTGATGTTT CCATTTTACA GACAGAAAAC
2901 TGAGTCTTAG ACAGAGGCCT AGAGGAGGCC AAGAGGTGGT GGGGCCAGGT
2951 CGGGGGGGCC CTGATGCTTG CTTCTCTCGC TTTGTTGTCAG CCCCAGGCA
3001 TGATGAAGAC TTTGTCCAGC GGGAACTGCA CGCTCAGTGT GCCCAGGAAA
3051 AACTCATACC GCATGGTGGT GCTGGGTGCC TCTCGGGTGG GCAAGAGCTC
3101 CATCGTGTCT CGCTTCTCTA ATGGCCGCTT TGAGGACCAG TACACACCCA

```

FIGURE 3, page 1 of 6

3151 CCATCGAGGA CTTCCACCGT AAGGTATACA ACATCCGCGG CGACATGTAC
 3201 CAGCTCGACA TCCTGGATAC CTCTGGCAAC CACCCCTTCC CCGCCATGCG
 3251 CAGGCTGTCC ATCCTCACAG GTGAGGCCCA CTGGTGCCTG GGCTGGGGCG
 3301 GCAGGGCCAG GGCATGGGTG CGGAGTGTGC TGGGCACTTG GCAGTTTGCA
 3351 TAGACTTGCA TAGCCATCGT CTGAGACAGG CGTCATCCCT GCACAATGAG
 3401 GCTCAGAGAG GTTTTGCCAT GTGCTGGAAA TAGTGATGAA GTCGGGGGCC
 3451 CCGATTCCAT TCTGTTAGAC TCCAGATCGA TTACTCATGG CTGTCGGGGC
 3501 CGCCTTCCAG ATCAGGAGCT GATACCAGCA TGCCCCAGGG ATATTCTTTT
 3551 CTAGGGAACA GAATGATGCC CTGGCTGCTG CTTTCCTTCT CCGGAAGATG
 3601 ACCCACCAGA GCTCCAGGGC CCAAGGTCAG TCCACGGGGC TCAGGTCTCC
 3651 CACACCCAG AGCCTTTGCCA CCTCCTAGAG AGGTAAGGGC AGGACCCAGG
 3701 CAGTGATCAC CAAAGGGAAG GGGGCTTGGT CATGGTCATA GTGATGGTGA
 3751 TGGCACTAGC TGACACTTAT CAGAAGCTAT GGGCCTGGCC CTGTTCTTAG
 3801 AGCTTGGCAT GTAGTTTTTT TTGAAACAGA GTCTCGCTCT GTACCCAGG
 3851 CTGGAGTGCA GTGGCGCGAT CTCGGCTCAC TGCAACCTCT GCCTCCCGGG
 3901 TTCAAGCGAT TCTCCTGCC T CAGTCCCCCA AGTAGCTGGG ACTACAGGCA
 3951 CGTGCCACCA TGCCCGGCTA ATTTTTTTGTA TTTTACTAG AGACGGGATT
 4001 TCACCATGTT AGCCAGGATG GTCTCGATCT CCTGACCTCG TGATCTGCCT
 4051 ACCTCAGCCT CCCAAAGTGC TGAGATTACA GGCCTCAGCC ACCGCGGCCA
 4101 GCCAGCATGT AGTTATTTAA CCCTCACAGT AAATAGTTAT TCATTCCCTT
 4151 TTTACAGGTG GGGAAACTGA AGCCAGAGA GGTAAAGTAA CTCACTCCAG
 4201 TGGTAGCACA GCTCGTAAAG GCAGTCTGCT TTTGTGCTTT CAGACAAAGC
 4251 CATACCACAG CCTCTCAGCC CTGCTGGGAA GGGTGAGGAG GGACAGGGAG
 4301 GTTGGGGGGA AGAAGGGGTG AGTGGAGCTG AGGGGCTGTG CCCTTGTTTTA
 4351 CACTGCATTA GCATGGTAGC TAAGAGGACA AGCCCGGGAC CCAGCACCTG
 4401 GGTGTGAGCC CTGGTTCCGC TGCTTCCTGG CTTTGTACCT CGAGGCAAGG
 4451 GATTTTATCT CTTGTGTCT CAGCACTCTC ATCTGTAGA CTGCACCACA
 4501 TCAACACTCA TCCTAAAGGG ACTGTGAGAC TTAAATGAAT GAATATATGT
 4551 AAGGCGCTTG GTGAGCAGAT AGTAAATGCA CAATAAATCC CCAAGTCTTC
 4601 TGTGTAGTCA GCATTTGCAA GTGGGCTGC TACGGGTAC ACGATCATTT
 4651 CCCAAGTCAC GCCCCTGAAG TTGCTGAGCA GGGATAAGGG AAGGAGTGAG
 4701 CAGGCAACTC TCTAGGCATC ATTCAGATAA CCCCCCACT GAGGTACTTC
 4751 TATACAGAGA AACCCATGCC ACTCCCAGCC CTGCTGCCGC CTTGAGGCCA
 4801 AGACTGAGGC TGCGGGGTGG CCCCTCCTTG AGTGCTTTCT CTTCCAGGCT
 4851 GGCTTTTCCG AGCATCTGAC CCAGACAGCA GTCAAGTTCT CCGCTCCACC
 4901 CCGAGTTTTC GAGAAGGGGC ATGTAGATAG GAGAGCCCTG GGTGACCTG
 4951 TGTTGGAATC CTTGCTGGGC CTCTTGGCTT AAATGTGTGA CCAGAGGCAC
 5001 ATGCATCCTG TCTGAGTCTC AGCCTCCCCA GCCACACAGT GGGCTTAACC
 5051 TCATACCCCG CAGGGAGGCT GTGAGGACTG CAAGAAGGCT TGTGGCGGGA
 5101 GCTTCCAGCA CGTGACGGGT ATTGCATTGG TGTGAGCTCC CCCAGCCTTG
 5151 GGGAGGGGAC TGGGTACCCG CTGCAATGAA TAAGGCTAAT GACAGAGGGA
 5201 AGGAGAGGGG AGATGTAGAG AGGAAGCACA TGCATATTTT CAGCATTAAT
 5251 TTTCACTGAC ACAAGTAATA CCAAACACA CCTCCTGCA AACGTACAG
 5301 ATAAAGCTAA TGCCCCTTTG ACCCATGTCC CCAATCCCAG GCTCCTGCCC
 5351 CTGCCCCGGA GGTGGCCACC CTGGCAGTCT GGCATGGAGC CTTCCGGGCC
 5401 TCCGTGACTA CACCGGCATT CGTATTTGTA TCCCCACAAT GGAGAGTATT
 5451 TTTGTCTGTC TCTTTTTTAT GCGCATATC ATTCTGAGCA CAGCTGTCTG
 5501 ATGCTTGTTT TTTTTTCACA CACCAACCCG TGCCTCATTT TCCAACCTGG
 5551 TGGAACCTCA TTTTTTCAAC CTCATTTTCC TGCTGCTCAG GAAATTCTGA
 5601 AAGCCATTAA TTCCACTGCC AGCTCTTCTT CCAGCTGCCA GACGGGCCGA
 5651 TCTCTGATGC TTGGCATCGC AGTCTCGCAT TTGAATATGT CAAGGCCACG
 5701 AGTCTCAGG GGGCCCGGAT TAGCCTAACG GGATGGGGTT TGGCAGCCCA
 5751 AGCAGGAAGA GTTGCCAAGC TGACGCTGCC TCGCAAGTGC CTTTCAGAAG
 5801 AGCCACACT GCAGTTCCTT CCATCACCTC CCATCCATTC AGGCTTCCTT
 5851 GGTAAACACT GACTGTGTGC CAGGCCCTGG GGAGACCAGG ACGAGTGGGT
 5901 GATGGAACCC TTCTCTGTGC CCGAGCTGTT TGGAGCACAC CTTTGATCTG
 5951 GACACCATTC TGAATGTGCC ATGTGCCATT AAATGGGGGT AAATGATGTG
 6001 CTCTGGGGGT GCAGAGGAAG GTGGCAGCCA TTCTGCCAGA AGCTGGAAC
 6051 GGTTGCTTCT CTTCTCAAGA ATTTGGGCCA ATTGCTGATT CCTCTGGGCC
 6101 TCAGTTTCTT CATCTGTGAG ACAGGGATCT TGTCACACCA CAAGGCTATC
 6151 AAGAGTTTGA GCAAAAGTGG TTGGACGCAG TGGCTCATGC CTGTAATCCC
 6201 AGCTCTTTGG GAGGCCGAGG TGGGCAGATC TCTTGAGGTC AGGAGTTCAA
 6251 GACCAGCCTG GCTAACACAG TGAAACACCG TCTCTACTAA AAAATACAAA

FIGURE 3, page 2 of 6

6301 AAATTAGCCA GGTGTGGTGA TGGGCACCTG TAATCCCAGT TACTCGGGAG
6351 GCTGAGGCAG GAGAATCTCT TGAACCCAGG AGGTGGAGGT TGCAGTGAGC
6401 TGAGATCTTG CCATTGCATT CCAGGCTGGG CAACAAGAGT GAAACTCTGT
6451 CTCAGAAAAA ATATAAATAA ATAAATAAAA AATAGCTAGG CATGGTGACA
6501 GCGCCTGTGA ATCCCAGCTG CTCCAGAGGC TGAGGCAGGA GAATCGCTTA
6551 AACCAGGAG GTGGAAGTTG CAGTGAGCCA AGATCACACC ACTGCACTCC
6601 AGCCTGGGCC ACAGAGAAAG ACTCCATCTC AAAAAAAAAA AAAAAAAAAA
6651 AAAAAGTTTA AGCAAAAGTG AGGAAGGTGC TTATTTAAAG CTGGAAATCA
6701 GGATGGAGGT ACCAGTCCAG ACAGCCTCCC CACCACCCCA CCGTCTCCAC
6751 AGCAGCCCCT GTTTCAGATT CACAAGCCTG CCTTGAGTGA TGCAGTGAGT
6801 TATCCTGGAG GCAGTGTGGG CCTTGGAGGC CAGCACTCAC TTTTTCATCC
6851 TATGATTTAT TTGAGAAGCA GAGAGCACCT ACCGGGTGCC AGGAACGAGC
6901 TAGGTGAGAA CAGAATCAGG TAGAAATCTC AGCCTAGCCA CACGGAAGCT
6951 GTGTGATCTT GGGCAGGCTG CATAACCCTT CTGAGCCTCA GTTTGCTCAC
7001 CTGTAATGCA AAGGTAACAA AATCTTGACA GAGGCATAGT GAGGAATCAA
7051 GAGAACAACG GGCCTGGAGC ATACACCCAG TGCTTAGCCC CCAGTAGGCC
7101 CTCACCTCTA TCATTACTGA CACCTGAGGT CACTGAGCAT GTGCCACTGT
7151 CCATTCACTA TCTTGCTATA CTCCCAAAAT CATCCTGCAA GGTAAATATT
7201 CATCTTCATG AAACAGACAG AGAAACTGAG GTTACAGAGG TTTCTGATC
7251 TGCCCAAGTC TGCTGGCAGC TAAGCGGATG AGGCCAGATG CAACTAGGC
7301 ATTGAGCAAG ACAGGCAGGA CCCCTGCTCT CATAGAAATG ATTTTATTA
7351 TTATCTGAAC ACAGTCCACA CAAGTGACCT ACCCCTCTCC AGCCCTGCAA
7401 AGAAATGTGA AGTGAGTTAA CTGTATTTGA ACCAAGTGGT CCACGTGTTA
7451 GCTATGCGAG TGTGAACAGG GGCTTCAACC CCCTCAGCCT CAGTTTCCTG
7501 TCCTGGAAAA TAATCGCAGG GAGAATAATC GCAGCTACCC CGAAGAGTCG
7551 CTGTGTAGGT TAAAGCAGTT ATGCCGCATA ACTGCTTCAG GGCACCTGTG
7601 ACTCCCAGCT CTTAGGGCTG ATGTTCTGTG GCCAGAGGAG GGCAGGGGTT
7651 GCAGTGGGCC GGTGAACTCA CTACCTGGGC TCTCTCCCTG CAGGGGATGT
7701 CTTCACTCTG GTGTTAGGCC TGGATAACCG GGAGTCCCTC GATGAGGTCA
7751 AGCGCCTTCA GAAGCAGATC CTGGAGGTCA AGTCTTGCCT GAAGAACAAG
7801 ACCAAGGAGG CGGCGGAGCT GCCCATGGTC ATCTGTGGCA ACAAGAACGA
7851 CCACGCGCAG CTGTGCGGCC AGGTGCCAC CACCAGGCC GAGCTGCTGG
7901 TGTCGGGCGA CGAGAACTGC GCCTACTTCG AGGTGTCGGC CAAGAAGAAC
7951 ACCAACGTGG ACAGATGTT CTACGTGCTC TTCAGCATGG CCAAGCTGCC
8001 ACACGAGATG AGCCCCGCCC TGCATCGCAA GATCTCCGTG CAGTACGGTG
8051 ACGCCTTCCA CCCCAGGCCC TTCTGCATGC GCCGCGTCAA GGAGATGGAC
8101 GCCTATGGCA TGGTCTCGCC CTTGCGCCGC CGCCCCAGCG TCAACAGTGA
8151 CCTCAAAGTAC ATCAAGGCCA AGGTCTTTCG GGAAGGCCAG GCCCCTGAGA
8201 GGGACAAGTG CACCATCCAG TGAGCGAGGG ATGCTGGGGC GGGGCTTGGC
8251 CAGTGCCTTC AGGGAGGTGG CCCCAGATGC CCACTGTGCG CATCTCCCA
8301 CCGAGGCCCG GGCAGCAGTC TTGTTACAG ACCTTAGGCA CCAGACTGGA
8351 GGCCCCCGGG CGCTGGCCTC CGCACATTCG TCTGCCTTCT CACAGCTTTC
8401 CTGAGTCCGC TTGTCCACAG CTCCTTGGTG GTTTCATCTC CTCTGTGGGA
8451 GGACACATCT CTGCGCCCTC AAGAGTTAGG CAGAGACTCA AGTTACACCT
8501 TCCTCTCCTG GGGTTGGAAG AAATGTTGAT GCCAGAGGGG TGAGGATTGC
8551 TGCGTCATAT GGAGCTCCT GGGACAAGCC TCAGGATGAA AAGGACACAG
8601 AAGGCCAGAT GAGAAAGGTC TCCTCTCTCC TGGCATAACA CCCAGCTTGG
8651 TTTGGGTGGC AGCTGGGAGA ACTTCTCTCC CAGCCCTGCA ACTCTTACGC
8701 TCTGGTTCAG CTGCCTCTGC ACCCCCTCCC ACCCCCAGCA CACACACAAG
8751 TTGGCCCCCA GCTGCGCCTG ACATTGAGCC AGTGGACTCT GTGTCTGAAG
8801 GGGGCGTGGC CACACCTCCT AGACCACGCC CACCACTTAG ACCACGCCCA
8851 CCTCCTGACC GCGTTCCTCA GCCTCCTCTC CTAGGTCCCT CCGCCCGACA
8901 GTTGTGCTTT GTTGTGGTTG CAGCTGTTTT CGTGTCATGT ATAGTAGTAG
8951 AAATGGAAAT CATTGTACTG TAAAAGCCTA GTGACTCCCT CCTTGGCCAG
9001 GCCCTCACCC AGTTCAGATC CACGGCCTCC ACCCGGGACG CCTTCTCCT
9051 CTGCTCCCAA ACAGGGTTTC CGTGGCCTGT TTGCAGCTAG ACATTGACCT
9101 CCGCCATTGA GCTCCAGGTT TACAGACAA TTGCACAAGC GTGGGGTGGG
9151 CAGGCCAGGA CTGCTTTTTT TTAATGCTCC CATTTACAG AGGATACCAC
9201 CGAGACTCGG AGGGGACACG ATGAGCACC AAGCCCCACT TTGTCCCCTA
9251 GCAAATTCAG GGTACAGCTC CACCTAGAAC CAGGCTGCCC TCTACTGTGC
9301 TCGTTCTCTA AGCATTATT AAGCACCTAC TGGGTGCTGG GTTCACTGTG
9351 TCCTAGGAAA CCAAGAGGGT CCCCAGTCTC GGCTCTGCC CGCCCTGCT
9401 GCCCCACCAC CTTCTGCACA CACAGCGGTG GGGAGGCGGG GAGGAGCAGC

FIGURE 3, page 3 of 6

```

9451 TGGGACCCAG AACTGAGCCT GGGAGGGATC CGACAGAAAA GCTCAGGGCG
9501 GGTCTTCTCC TTGTGCCCCG GATTGGGCTA TGCTGGGTAC CACCATGTAC
9551 TCAGGCATGG TGGGTTTTGA ACCCATAAAC CAAAGGCCCT TGTATCAGC
9601 TCTTAACAAG TATATTTTGT ATTTTAATCT CTCTAAACAT ATTGAAGTTT
9651 TAGGGCCCTA AGGAACCTTA GTGATCTTCT ATTGGGTCTT TCTGAGGTTC
9701 AGAGAGGGTA AGTAACTTCC TCCAGGTCAC ACAGCAAGTC TGTGGGTGGC
9751 AGAAGCAAGC TAGCGCTGGG CATTGAGTAC ATACCACGAT GTGCTCCCTC
9801 TCTTGATGCT TGGCCCCCTG GGCCTTCAGG GCTTTGGGAC ATCTTGTCCT
9851 CAACCTCTC CCTAGATCAG TCTGTGAGGG TCCCTGTAGA TATTGTGTAC
9901 ACCATGCCCC TGTATATACA AGTACACACA GATGTACACA CAGATGTACA
9951 CATGCTCCAG CCCAGCTCT GCATACCTGC ACCTGCACCC CAGCCTTGGC
10001 CCCTGCCTGC GTCTGTGCTC AAAGCAGCAG CTCCAACCCT GCCTCTGTCC
10051 CCTTCCCCAC CCACTGCCTG AGCCTTCTGA GCAGACCAGG TACCTTGGCT
10101 GCACCGGTGT GTGGCCCGCT CTCACCCAGG CACAGCCCCG CCACCATGGA
10151 TCTCGTGTA CACTATCAAT AAAAGTGGGT TTGTTACAAA GCCGTGTCCT
10201 TGCCCATGTG TATTTTTTGT ATTTCCAAGA GGAGGTGTGC CCCTTCCAG
10251 ACCAAAGCTG GCCTTTCCT CCCAAAATGC ACCTGCCGTG TACCCTGGCC
10301 CTGAGGGTCA GCACTGAGTC CACCTTCAAG TGTAAGTGTG GGGAGAGGGG
10351 GATAAGTCCC CCAGATGGAA GGTGATGCCC TCCTTCAGCC TGGCCCTCCT
10401 GGGTCCTCCG GGTGTGTGTA CCGAGGTGTC TGTGTCCACA AAGAAGGGGC
10451 CCCCCTGGAC CATTAGCTCC AGGAGGATCT CCGTGTCTGA GTTCTTTGTG
10501 ATTCTGTGTA AGCAGCAATT TCACCCGCAG GGGACAGTTG GCAATCTCTG
10551 GAAACCTTTT CCAAGCCTGG GGCTGGGGCT GCTACTCTCA TCTGGTGGGT
10601 GGAGGCCAGG GACACCATTG AGTATCCTCC AACGCACAGG ATGCCCTCC
10651 ACCCCACCCC CACTGAGAAT TATCTGGCCT CAAATGCCAA GCGTGGGCAG
10701 CCTTACTTAG ACTCACCCCA GGGGCTGGGA CACGCCCCCA CCTGCGTGTG
10751 ATGGATTTGT TGGACCACAT TCTGGACGGA ACCCACAGCA TAAGCACTCC
10801 TGTGAAGTGA GACAGGATGT GGGTGAGGAT GGAAAGTGGA GGCTGAGGGA
10851 GAAGGTCTGG GCCCTGACCA ACACGGAATG TGCCCCCTGG GACTGAGAGG
10901 CTTCCCTGGG CAGAGGGAAA GGAGGAAGTC AGTGAGGTAA AATACTCCCT
10951 GTGTGTTTTA CCCAGCGAGT CTCACGCCAT CCTATCACCC AGCCCCGAGG
11001 GAAGCCCACT CATGTTCAAC CCATCTGAGC ATTTAGGCTC AGAGAGCTCA
11051 ATATCTTGTC CAAGATGGCA CAGCTGGTGA AGTGGCAGAT CAGAGATTCA
11101 ACACCAGAGG CTGTCTGATT TCCGTCTGGC TGAAGAAAGA TTTTGCATCA
11151 GGGAGGTGGA AACCATCTGT GCTTTTGATC AGCAAATGCC ACCAGCAGGA
11201 TCAGGGAGCC AGGCCATAAA G

```

FEATURES:

```

Start: 3000
Exon: 3000-3270
Intron: 3271-7693
Exon: 7694-8220
Stop: 8221

```

CHROMOSOME MAP POSITION:

Chromosome 22

ALLELIC VARIANTS (SNPs):

DNA			
Position	Major	Minor	Domain
3951	C	T	Intron
4127	C	T	Intron
4157	G	A	Intron
4513	C	T	Intron
6894	C	A	Intron
8409	G	C	Beyond ORF (3')
8437	T	G	Beyond ORF (3')
8579	T	C	Beyond ORF (3')
10292	A	G	Beyond ORF (3')
10792	A	G	Beyond ORF (3')

Context:

DNA
Position
3951

CACACCCCAGGCCTTTGCCACCTCCTAGAGAGGTAAGGGCAGGACCCAGGCAGTGATCAC
CAAAGGGAAGGGGGCTTGGTCATGGTCATAGTGATGGTGATGGCACTAGCTGACACTTAT
CAGAAGCTATGGGCTGGCCCTGTTCTTAGAGCTTGGCATGTAGTTTTTTTTTGAACAGA
GTCTCGCTCTGTACCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAACCTCT
GCCTCCCGGTTCAAGCGATTCTCCTGCCTCAGTCCCCAAGTAGCTGGGACTACAGGCA
[C, T]
GTGCCACCATGCCCCGGCTAATTTTTGTATTTTTTACTAGAGACGGGATTTACCATGTTA
GGCAGGATGGTCTCGATCTCCTGACCTCGTGATCTGCCTACCTCAGCCTCCCAAAGTGCT
GAGATTACAGGCGTCAGCCACCGCGGCCAGCCAGCATGTAGTTATTTAACCTCACAGTA
AATAGTTATTTCATTCCCTTTTTACAGGTGGGGAACTGAAGCCCAGAGAGGTTAAGTAAC
TCACTCCAGTGGTAGCACAGCTCGTAAAGGCAGTCTGCTTTTGTGCTTTTCAGACAAAGCC

4127 CAGAGTCTCGCTCTGTACCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAC
CTCTGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGTCCCCAAGTAGCTGGGACTACA
GGCAGGTGCCACCATGCCCGGCTAATTTTTGTATTTTTTACTAGAGACGGGATTTACCA
TGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCTGCCTACCTCAGCCTCCCAA
GTGCTGAGATTACAGGCGTCAGCCACCGCGGCCAGCCAGCATGTAGTTATTTAACCTCA
[C, T]
AGTAAATAGTTATTTCATTCCCTTTTTACAGGTGGGGAACTGAAGCCCAGAGAGGTTAAG
TAACTCACTCCAGTGGTAGCACAGCTCGTAAAGGCAGTCTGCTTTTGTGCTTTTCAGACA
AGCCATACCACAGCCTCTCAGCCTCCCAAAGTGCTGAGATTACAGGCGTCAGCCACCGCG
GGAAGAAGGGGTGAGTGGAGCTGAGGGGCTGTGCCCTTGTTTACACTGCATTAGCATGGT
AGCTAAGAGGACAAGCCCGGACCCAGCACCTGGGTGTGAGCCCTGGTTCGCTGCTTCC

4157 TGCAGTGGCGCGATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCT
GCCTCAGTCCCCAAGTAGCTGGGACTACAGGCACGTGCCACCATGCCCGGCTAATTTTT
TGATTTTTTACTAGAGACGGGATTTACCATGTTAGCCAGGATGGTCTCGATCTCCTGAC
CTCGTGATCTGCCTACCTCAGCCTCCCAAAGTGCTGAGATTACAGGCGTCAGCCACCGCG
GCCAGCCAGCATGTAGTTATTTAACCTCACAGTAAATAGTTATTTCATTCCCTTTTTACA
[G, A]
GTGGGGAACTGAAGCCCAGAGAGGTTAAGTAACTCACTCCAGTGGTAGCACAGCTCGTA
AAGGCAGTCTGCTTTTGTGCTTTTCAGACAAAGCCATACCACAGCCTCTCAGCCCTGCTGG
GAAGGGTGAGGAGGGACAGGGAGGTTGGGGGGAAGAAGGGGTGAGTGGAGCTGAGGGGCT
GTGCCCTTGTTTACACTGCATTAGCATGGTAGCTAAGAGGACAAGCCCGGACCCAGCAC
CTGGGTGTGAGCCCTGGTTCGCTGCTTCTGGCTTTGTACCTCGAGGCAAGGGATTTTA

4513 TCGTAAAGGCAGTCTGCTTTTGTGCTTTTCAGACAAAGCCATACCACAGCCTCTCAGCCCT
GCTGGGAAGGGTGAGGAGGGACAGGGAGGTTGGGGGGAAGAAGGGGTGAGTGGAGCTGAG
GGGCTGTGCCCTTGTTTACACTGCATTAGCATGGTAGCTAAGAGGACAAGCCCGGGACCC
AGCACCTGGGTGTGAGCCCTGGTTCGCTGCTTCTGGCTTTGTACCTCGAGGCAAGGGA
TTTTATCTCCTTGTGTCTCAGCACTCTCATCTGTAAGACTGCACCACATCAACACTCATC
[C, T]
TAAAGGGACTGTGAGACTTAAATGAATGAATATATGTAAGGCGCTTGGTGAGCAGATAGT
AAATGCACAATAAATCCCCAAGTCTTCTGTGAGTCAAGTATTTGCAAGTGGGCCTGCTAC
GGGTTACACGATCATTTCCCAAGTCACGCCCCTGAAGTTGCTGAGCAGGGATAAGGGAAG
GAGTGAGCAGGCAACTCTTAGGCATCATTAGATAACCCCCCACTGAGGTACTTCTAT
ACAGAGAAACCATGCCACTCCAGCCCTGCTGCCGCTTGAGGCCAAGACTGAGGCTGC

6894 GCACTCCAGCCTGGGCCACAGAGAAAGACTCCATCTCAAAAAAAAAAAAAAAAAAAAAA
AAGTTTAAGCAAAAGTGAGGAAGGTGCTTATTAAGCTGGAAATCAGGATGGAGGTACC
AGTCCAGACAGCCTCCCCACCACCCACCGTCTCCACAGCAGCCCCTGTTTCAGATTAC
AAGCCTGCCTTGAGTGATGCAGTGAGTTATCCTGGAGGCACTGTGGGCCTTGAGGCCAG
CACTCACTTTTTCATCTATGATTATTTGAGAAGCAGAGAGCACCTACCGGCTGCCAGG
[C, A]
ACGAGCTAGGTGAGAACAGAATCAGGTAGAAATCTCAGCCTAGCCACACGGAAGCTGTGT
GATCTTGGGCAGGCTGCATACCCTTTCTGAGCCTCAGTTTGCTCACCTGTAATGCAAAGG
TAACAAAATCTTGACAGAGGCATAGTGAGGAATCAAGAGAACACGGGCCTGGAGCATAC
ACCCAGTGCTTAGCCCCAGTAGGCCCTCACTCTCATCATTACTGACACCTGAGGTCACT
GAGCATGTGCCACTGTCCATTCAATATCTTGATAACTCCCAAAATCATCTGCAAGGTA

FIGURE 3, page 5 of 6

8409 CATGGTCTCGCCCTTCGCCCCGCCGCCAGCGTCAACAGTGACCTCAAGTACATCAAGGC
CAAGGTCTTCGGGAAGGCCAGGCCGTGAGAGGGACAAGTGCAACCATCCAGTGAGCGAG
GGATGCTGGGGCGGGGCTTGGCCAGTGCCTTCAGGGAGGTGGCCCCAGATGCCCACTGTG
CGCATCTCCCCACCGAGGCCCGGCCAGCAGTCTTGTTCACAGACCTTAGGCACCAGACTG
GAGGCCCCCGGGCGCTGGCCTCCGCACATTCTGTCTGCTTCTCACAGCTTTCCTGAGTCC
[G, C]
CTTGTCCACAGCTCCTTGGTGGTTTCATCTCCTCTGTGGGAGGACACATCTCTGCAGCCT
CAAGAGTTAGGCAGAGACTCAAGTTACACCTTCCTCTCCTGGGGTTGGAAGAAATGTTGA
TGCCAGAGGGGTGAGGATTGCTGCGTCATATGGAGCCTCCTGGGACAAGCCTCAGGATGA
AAAGGACACAGAAGGCCAGATGAGAAAGGTCTCCTCTCTCCTGGCATAACACCCAGCTTG
GTTTGGGTGGCAGCTGGGAGAACTTCTCTCCAGCCCTGCAACTCTTACGCTCTGGTTCA

8437 AGCGTCAACAGTGACCTCAAGTACATCAAGGCCAAGGTCTTCGGGAAGGCCAGGCCCGT
GAGAGGGACAAGTGCAACCATCCAGTGAGCGAGGGATGCTGGGGCGGGGCTTGGCCAGTGC
CTTCAGGGAGGTGGCCCCAGATGCCCACTGTGCGCATCTCCCCACCGAGGCCCGGCCAGC
AGTCTTGTTCACAGACCTTAGGCACCAGACTGGAGGCCCGGGCGCTGGCCTCCGCACA
TTCGTCTGCCCTTCTCACAGCTTTCCTGAGTCCGCTTGTCCACAGCTCCTTGGTGGTTTCA
[T, G]
CTCCTCTGTGGGAGGACACATCTCTGCAGCCTCAAGAGTTAGGCAGAGACTCAAGTTACA
CCTTCCTCTCCTGGGGTTGGAAGAAATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTCA
TATGGAGCCTCCTGGGACAAGCCTCAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAG
GTCTCCTCTCTCCTGGCATAACACCCAGCTTGGTTTGGGTGGCAGCTGGGAGAACTTCTC
TCCCAGCCCTGCAACTCTTACGCTCTGGTTTCACTGCCTCTGCACCCCTCCACCCCCA

8579 GCCCACTGTGCGCATCTCCCCACCGAGGCCCGGCCAGCAGTCTTGTTCACAGACCTTAGG
CACCAGACTGGAGGCCCGGGCGCTGGCCTCCGCACATTCTGTCTGCCTTCTCACAGCTT
TCCTGAGTCCGCTTGTCCACAGCTCCTTGGTGGTTTCATCTCCTCTGTGGGAGGACACAT
CTCTGCAGCCTCAAGAGTTAGGCAGAGACTCAAGTTACACCTTCCTCTCCTGGGGTTGGA
AGAAATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTCATATGGAGCCTCCTGGGACAAG
[T, C]
CTCAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAGGTCTCCTCTCTCCTGGCATAAC
ACCCAGCTTGGTTTGGGTGGCAGCTGGGAGAACTTCTCTCCAGCCCTGCAACTCTTACG
CTCTGGTTTCACTGCTCTGCAACCCCTCCACCCCGACACACACAAGTTGGCCCCC
AGCTGCGCCTGACATTGAGCCAGTGGACTCTGTGTCTGAAGGGGGCGTGGCCACACCTCC
TAGACCACGCCACCACTTAGACCACGCCACCTCTGACCGCTTCTCAGCCTCCTCT

10292 AGCCTTGGCCCCCTGCCTGCGTCTGTGCTCAAAGCAGCAGCTCCAACCCCTGCCTCTGTCCC
CTTCCCCACCACTGCCTGAGCCTTCTGAGCAGACCAGGTACCTTGGCTGCACCGGTGTG
TGGCCCGCTCTCACCCAGGCACAGCCCCGCCACCATGGATCTCCGTGTACACTATCAATA
AAAGTGGGTTTGTACAAAGCCGTGTCTTGGCCATGTGTATTTTTTGTATTTCCAAGAG
GAGGTGTGCCCCCTTCCAGACCAAAGCTGGCCTTTCCTCCCAAATGCACCTGCCGTGT
[A, G]
CCCTGGCCCTGAGGGTCACTGAGTCCACCTTCAAGTGTAAGTGTGGGGAGAGGGGGA
TAAGTCCCCCAGATGGAAGGTGATGCCCTCCTTCAGCCTGGCCCTCCTGGGTCTCCGGG
TGTGTGTACCGAGGTGTCTGTGTCCACAAAGAAGGGGGCCCCGTGGACCATTAGCTCCAG
GAGGATCTCCGTGTCTGAGTTCTTGTGATTCTGTACAGCAGCAATTCACCCGCAGGG
GACAGTTGGCAATCTCTGGAACCTTTTCCAAGCCTGGGGCTGGGGCTGCTACTCTCATC

10792 TTCTTTGTGATTCTCTGTACAGCAGCAATTTACCCGCAGGGGACAGTTGGCAATCTCTGG
AAACCTTTTCCAAGCCTGGGGCTGGGGCTGCTACTCTCATCTGGTGGGTGGAGGCCAGGG
ACACCATTAGTATCCTCCAACGCACAGGATGCCCCCTCCACCCCACTGAGAAAT
ATCTGGCCTCAAATGCCAAGCGTGGGCAGCCTTACTTAGACTCACCCAGGGGCTGGGAC
ACGCCCCACCTGCGTGTGATGGATTTGTTGGACCACATCTGGACGGAACCCACAGCAT
[A, G]
AGCACTCCTGTGAAGTGAGACAGGATGTGGGTGAGGATGGAAAGTGGAGGCTGAGGGAGA
AGGTCTGGGGCCTGACCAACACGGAATGTGCCCTTGGGACTGAGAGGCTTCCCTGGGCA
GAGGGAAAGGAGGAAGTCAGTGAGGTAAAATACTCCCTGTGTGTTTTACCCAGCGAGTCT
CACGCCATCCTATCACCCAGCCCCGAGGGAAGCCCACTCATGTTACCCCATCTGAGCAT
TTAGGCTCAGAGAGCTCAATATCTTGTCCAAGATGGCACAGCTGGTGAAGTGGCAGATCA

FIGURE 3, page 6 of 6